AD-A105 745 FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS F/G 12/1
NEW ESTIMATION METHODS FOR LOG-LINEAR MODELS.(U)
AUG 81 T C REDMAN
UNCLASSIFIED FSU-STATISTICS-M562 N00014-76-C-0394
NL

1 OF 1
AD
A105745

END
DATE
FILMED
11 81
DTIC

LEVEL II

12

AD A105745

The Florida State University

Department of Statistics

Tallahassee, Florida
32306

NEW ESTIMATION METHODS FOR LOG-LINEAR MODELS.

by

Thomas C. Redman

FSU Statistics Report No. M562
ONR Technical Report No. 151

~~August,~~ 1981

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

Research supported in part by the Office of Naval Research under contracts N00014-76-C-0394 ~~and~~ N00014-80-C-0093 and in part by the National Institute ~~of Environmental Health Science in the~~ Grant 5 T32 ES07011. Reproduction in which or in part is permitted for any purpose of the United States Government.

NEW ESTIMATION METHODS FOR LOG-LINEAR MODELS[1]

by

Thomas C. Redman[2]

The Florida State University

Tallahassee, Florida 32306

Two new methods for estimation of parameters in log-linear models are proposed and their properties considered in this article. I~~n particular,~~ Conditions for the existence of the new estimators~~, more easily checked than those available for likelihood estimators,~~ are derived, and the new estimators are shown to possess appropriate asymptotic properties. KEY WORDS: Minimum δ estimation, approximate minimum δ estimation, existence, limiting distribution.

---

1. Research supported by the office of Naval Research under Contracts N00014-76-C-0394 and N00014-80-C-0093 and in part by the National Institute of Environmental Health Sciences under Grant S T32 ES07011. Reproduction in whole or in part is permitted for any purpose of the United States government.

2. Now a Member of Technical Staff, Bell Laboratories, Holmdel, N. J., 07733.

1. INTRODUCTION

Two new estimation methods for parameters in log-linear models are introduced in this paper. One of these particularly provides easy calculation and both are shown to have appropriate asymptotic properties. Existence results yield more easily verified conditions. The new methods are designated as the minimum d and the approximate minimum d method. We proceed with the development of needed notation and background.

Following Haberman (1974), we regard a contingency table as a k-vector $\underset{\sim}{n}$, $\underset{\sim}{n}' = (n_1, \ldots, n_k)$, where n_i denotes the number of observations in cell i from one of L multinomial distributions. Specifically, the sets $\{n_i: i \in I_\ell\}$, $\ell = 1, \ldots, L$, define independent multinomial vectors,

$$\{n_i: i \in I_\ell\} \sim M(N_\ell, \pi_i: i \in I_\ell, \pi_i > 0, \sum_{i \in I_\ell} \pi_i = 1),$$

where $N_\ell = \sum_{i \in I_\ell} n_i$, and $I_1, \ldots, I_L$ are disjoint and exhaustive subsets of $\{1, \ldots, k\}$. Let $\underset{\sim}{\pi}' = (\pi_1, \ldots, \pi_k)$ and $\underset{\sim}{p}' = (p_1, \ldots, p_k)$, where $p_i = n_i/N_\ell$ for $i \in I_\ell$. Corresponding to $\underset{\sim}{n}$ and $\underset{\sim}{p}$, let $\underset{\sim}{\gamma}' = (\gamma_1, \ldots, \gamma_k)$, where $\gamma_i = \log \pi_i$, $i = 1, \ldots, k$, and $\log \underset{\sim}{p}' = (\log p_1, \ldots, \log p_k)$. Finally, for a k-vector $\underset{\sim}{x}$, let $\exp \underset{\sim}{x}' = (\exp x_1, \ldots, \exp x_k)$.

The development of log-linear models was heralded by Bartlett (1935) and by Roy and Kastenbaum (1956). They define interactions in terms of constraints on products of elements of $\underset{\sim}{\pi}$. Birch (1963) shows that these constraints coincide with constraints on linear combinations of elements of $\underset{\sim}{\gamma}$. Most generally, in a log-linear model, the parameters $\underset{\sim}{\gamma}$ are constrained by $\underset{\sim}{B}_1\underset{\sim}{\gamma} = \underset{\sim}{0}_m$, $0 \leq m < k$; when $m=0$, there are no linear constraints on $\underset{\sim}{\gamma}$. The matrix $\underset{\sim}{B}_1$ is assumed to be an m x k matrix with orthonormal rows and to satisfy

$$\underset{\sim}{B}_1\underset{\sim}{A} = \underset{\sim}{0}_{m \times L}, \tag{1.1}$$

where $\underset{\sim}{A}$ is the k x L matrix given by

$$a_{i\ell} = \begin{cases} 1, & i \in I_\ell, \\ 0, & \text{otherwise}. \end{cases} \tag{1.2}$$

We denote the parameter space of $\underset{\sim}{\gamma}$ by $\Gamma(\underset{\sim}{B}_1)$. Let

$$\Gamma_S = \{\underset{\sim}{\gamma}: \sum_{i \in I_\ell} \exp \gamma_i = 1, \ \ell = 1, \ldots, L\}, \tag{1.3}$$

and

$$\Gamma_M(\underset{\sim}{B}_1) = \{\underset{\sim}{\gamma}: \underset{\sim}{B}_1\underset{\sim}{\gamma} = \underset{\sim}{0}_m, \ \underset{\sim}{\gamma} \in E^k\}. \tag{1.4}$$

Evidently

$$\Gamma(\underset{\sim}{B}_1) = \Gamma_S \cap \Gamma_M(\underset{\sim}{B}_1). \tag{1.5}$$

Since its rows are orthonormal, we may augment $\underset{\sim}{B}_1$ to a k x k orthonormal matrix through addition of additional rows $\underset{\sim}{B}_2$ and define

$$\underset{\sim}{B}' = [\underset{\sim}{B}_1', \underset{\sim}{B}_2']. \tag{1.6}$$

Reparameterization of $\underset{\sim}{\gamma}$ defines

$$\underset{\sim}{\mu} = \underset{\sim}{B}\, \underset{\sim}{\gamma}, \tag{1.7}$$

$$\underset{\sim}{\gamma} = \underset{\sim}{B}'\underset{\sim}{\mu}. \tag{1.8}$$

Since $\underset{\sim}{\gamma} \in \Gamma_M(\underset{\sim}{B}_1)$ implies $\underset{\sim}{B}_1\, \underset{\sim}{\gamma} = \underset{\sim}{0}_m$,

$$\underset{\sim}{\mu}' = (\underset{\sim}{\mu}_1', \underset{\sim}{\mu}_2') = (\underset{\sim}{0}_m', \underset{\sim}{\mu}_2'), \tag{1.9}$$

(1.8) reduces to

$$\underset{\sim}{\gamma} = \underset{\sim}{B}_2\, \underset{\sim}{\mu}_2', \tag{1.10}$$

and (1.10) provides an alternative formulation of the general log-linear model.

There are many methods for estimation of the parameters of a log-linear model, the method of maximum likelihood being the best known. Estimation equations for the general log-linear model are easily derived by the methods of Birch (1963) or LaGrange (Apostol, 1957), and, for L=1, are given by

$$\underset{\sim}{B}_1\, \underset{\sim}{\gamma} = \underset{\sim}{0}_m,$$

$$\underset{\sim}{B}_2\, [\underset{\sim}{p} - \exp \underset{\sim}{\gamma}] = \underset{\sim}{0}_{k-m}. \tag{1.11}$$

Generally, equations (1.11) require iterative solution. Conditions under which solutions may be obtained explicitly are available (Andersen, 1974, Bishop, Feinberg, and Holland, 1975).

Haberman (1973, 1974) and Andersen (1974) consider the existence and uniqueness of the likelihood estimator. Their results, in our notation, are summarized in the following theorem.

<u>Theorem 1.1</u>: If a maximum likelihood estimator $\hat{\gamma}$ of γ exists, it is unique, and satisfies (1.11). Furthermore, any of the following conditions is necessary and sufficient for the existence of a likelihood estimator:

i) there exists a vector $\hat{\gamma}$ which satisfies (1.11),

ii) there exists a vector ν such that $B_2\nu = 0_{k-m}$, and $p_i + \nu_i > 0$, $i = 1, \ldots, k$, and

iii) there exists no vector ν such that $B_1\nu = 0_m$, $\nu_i \leq 0$, $i = 1, \ldots, k$, $\nu \neq 0_k$, and $p'\nu = 0$.

Except in specific cases, the conditions of Theorem 1.1 are difficult to apply in practice.

The parameters of a log-linear model may be estimated also by the principle of weighted least squares (Grizzle, Starmer, and Koch, 1969, Koch <u>et al</u>. 1977).

Both the weighted least squares and the maximum likelihood approaches lead to estimators with good asymptotic properties. In Section 2, we introduce the two new estimation methods. Existence and uniqueness of the estimators are discussed in Section 3 and, in Section 4, it is shown that the new estimates also have good asymptotic properties.

2. MINIMUM δ ESTIMATION

2.1 The Minimum δ Method

Minimum δ estimation is not really new, as δ is included in the class of functions studied by Neyman (1949), Taylor (1953), and Ferguson (1958), the minimization of which leads to B.A.N. estimators. It does, however, seem to have special efficacy for log-linear models and leads to the approximate minimum δ estimation method discussed in Section 2.2.

Let

$$\begin{aligned} \delta(\underset{\sim}{\gamma};\ \underset{\sim}{n}) &= \sum_{i=1}^{k} n_i(\log p_i - \gamma_i)^2 \\ &= (\log \underset{\sim}{p} - \underset{\sim}{\gamma})'\underset{\sim}{N}\ (\log \underset{\sim}{p} - \underset{\sim}{\gamma}), \end{aligned} \tag{2.1}$$

where $\underset{\sim}{N}$ is the diagonal matrix with entries $n_1, \ldots, n_k$. If, for some i, n_i is zero, the contribution to $\delta(\underset{\sim}{\gamma};\ \underset{\sim}{n})$ of the i^{th} cell is taken to be zero. Further, we use

$$0\ (\log\ 0)^q = 0\ , \tag{2.2}$$

since

$$\begin{aligned} \lim_{n_i \to 0} n_i(\log p_i)^q &= \lim_{p_i \to 0} N_\ell [p_i(\log p_i)^q] \\ &= 0;\ i \in I_\ell,\ \ell = 1, \ldots, L. \end{aligned}$$

The minimum δ estimator $\tilde{\underset{\sim}{\gamma}}$ of $\underset{\sim}{\gamma}$ is defined to be that point in $\Gamma(\underset{\sim}{B}_1)$ which minimizes $\delta(\underset{\sim}{\gamma};\ \underset{\sim}{n})$. Equivalently, $\tilde{\underset{\sim}{\mu}}$ and $\tilde{\underset{\sim}{\pi}}$ denote the corresponding estimators of $\underset{\sim}{\mu}$ and $\underset{\sim}{\pi}$ respectively.

In analogy to the likelihood estimator, the minimum δ estimator satisfies a system of equation whenever it exists. This system is given by the following theorem.

Theorem 2.1: If a minimum δ estimator $\tilde{\gamma}$ of γ exists, then it satisfies the equations,

$$B_2[N(\log p - \tilde{\gamma}) - y(\tilde{\gamma})] = 0_{k-m}, \tag{2.3}$$

$$B_1\tilde{\gamma} = 0_m, \tag{2.4}$$

and

$$\sum_{i \in I_\ell} \exp \tilde{\gamma}_i = 1, \ \ell = 1, \ldots, L, \tag{2.5}$$

where $y(\gamma)' = [y_1(\gamma), \ldots, y_k(\gamma)]$,

$$y_j(\gamma) = (\exp \gamma_j) \sum_{i \in I_\ell} n_i(\log p_i - \gamma_i) \tag{2.6}$$

for $j \in I_\ell$, $\ell = 1, \ldots, L$.

Proof: The theorem will follow from LaGrange's Theorem (Apostol, 1957, Th. 7-10; Williamson et al., 1972, p. 595) upon verification of the conditions of that theorem. The set $\{\gamma: \gamma_i < 0, i = 1, \ldots, k\}$ is associated with the open set of the theorem. The constraint functions are

$$B_1\gamma,$$

and

$$\sum_{i \in I_\ell} \exp \gamma_i - 1, \ \ell = 1, \ldots, L,$$

and they vanish on $\Gamma(\underset{\sim}{B}_1)$.

The matrix

$$[\underset{\sim}{B}_1', \underset{\sim}{e}_1, \ldots, \underset{\sim}{e}_L]', \tag{2.7}$$

where $\underset{\sim}{e}_\ell' = (e_{\ell 1}, \ldots, e_{\ell k})$,

$$e_{\ell i} = \begin{cases} \exp \gamma_i, & i \in I_\ell, \\ 0, & i \in I_\ell, \end{cases} \tag{2.8}$$

is of full rank, since $\underset{\sim}{B}_1$ is of full rank and it may be proven that no row $\underset{\sim}{e}_\ell'$ is a linear combination of rows of $\underset{\sim}{B}_1$. The proof follows by contradiction. Suppose that

$$\underset{\sim}{e}_\ell' = \sum_{j=1}^{m} \alpha_i \underset{\sim}{B}_{1j},$$

where $\underset{\sim}{B}_{1j}$

denotes the j^{th} row of $\underset{\sim}{B}_1$. Then we know that

$$\underset{\sim}{e}_\ell' \underset{\sim}{1}_k = 1,$$

but

$$\underset{\sim}{e}_\ell' \underset{\sim}{1}_k = (\sum_{j=1}^{m} \alpha_j \underset{\sim}{B}_{1j}) \underset{\sim}{1}_k = \sum_{j=1}^{m} \alpha_j (\underset{\sim}{B}_{1j} \underset{\sim}{1}_k) = 0$$

by (1.3) and the contradiction

is established.

Finally, since it is apparent that d and the constraint functions have continuous partial derivatives with respect to γ_j, $j = 1, \ldots, k$, on $\{\underset{\sim}{\gamma}: \gamma_i < 0, i = 1, \ldots, k\}$, the conditions of LaGrange's Theorem are met.

Let

$$\psi(\underset{\sim}{\gamma}; \underset{\sim}{n}) = d(\underset{\sim}{\gamma}; \underset{\sim}{n}) + \underset{\sim}{\phi}'\underset{\sim}{B}_1\underset{\sim}{\gamma} + \sum_{\ell=1}^{L} \Delta_\ell \Big(\sum_{i \in I_\ell} \exp \gamma_i - 1\Big)$$

be the Lagrangian function for the minimization of d subject to $\underset{\sim}{\gamma} \in \Gamma(\underset{\sim}{B}_1)$. Then LaGrange's Theorem implies that, given $\underset{\sim}{\tilde{\gamma}} \in \Gamma(\underset{\sim}{B}_1)$, values of the m-vector $\underset{\sim}{\phi}$, and $\Delta_1, \ldots, \Delta_L$ exist satisfying the equations,

$$\left.\frac{d\, \psi(\underset{\sim}{\gamma};\, n)}{d\, \gamma_j}\right|_{\underset{\sim}{\gamma}=\underset{\sim}{\tilde{\gamma}}} = -2n_j(\log p_j - \tilde{\gamma}_j) + \sum_{i=1}^{m} \phi_i\, b_{1ij}$$

$$+ \Delta_\ell \exp \gamma_j = 0,\; j \in I_\ell,\; \ell = 1, \ldots, L, \tag{2.9}$$

where b_{1ij} is the (i,j) element of $\underset{\sim}{B}_1$. Summation of the subset of equations (2.9) over $j' \in I_\ell$ yields

$$-2 \sum_{j' \in I_\ell} n_{j'}(\log p_{j'} - \tilde{\gamma}_{j'}) + \Delta_\ell = 0, \tag{2.10}$$

the term involving b_{1ij}'s vanishing due to (1.1). With the use of (2.10) and vector notation, (2.9) becomes

$$-2\ \underset{\sim}{N}(\log \underset{\sim}{p} - \underset{\sim}{\bar{\gamma}}) + \underset{\sim}{B}_1'\underset{\sim}{\phi} + 2\underset{\sim}{y}(\underset{\sim}{\bar{\gamma}}) = \underset{\sim}{0}_k, \qquad (2.11)$$

where $\underset{\sim}{y}(\underset{\sim}{\bar{\gamma}})$ is given by (2.6). Premultiplication of (2.11) by $\underset{\sim}{B}_1$ yields

$$-2\ \underset{\sim}{B}_1\underset{\sim}{N}(\log \underset{\sim}{p} - \underset{\sim}{\bar{\gamma}}) + \underset{\sim}{\phi} + 2\ \underset{\sim}{B}_1\ \underset{\sim}{y}(\underset{\sim}{\bar{\gamma}}) = \underset{\sim}{0}_m, \qquad (2.12)$$

which implies

$$\underset{\sim}{\phi} = 2\ \underset{\sim}{B}_1\underset{\sim}{N}(\log \underset{\sim}{p} - \underset{\sim}{\bar{\gamma}}) - 2\ \underset{\sim}{B}_1\ \underset{\sim}{y}(\underset{\sim}{\bar{\gamma}}). \qquad (2.13)$$

Substitution of (2.13) into (2.11) yields

$$[\underset{\sim}{I} - \underset{\sim}{B}_1'\underset{\sim}{B}_1][\underset{\sim}{N}(\log \underset{\sim}{p} - \underset{\sim}{\bar{\gamma}}) - \underset{\sim}{y}(\underset{\sim}{\bar{\gamma}})] = \underset{\sim}{0}_k, \qquad (2.14)$$

and, with premultiplication by $\underset{\sim}{B} = \left|\begin{matrix}\underset{\sim}{B}_1 \\ \underset{\sim}{B}_2\end{matrix}\right|$ (2.14) reduces to

$$\underset{\sim}{B}_2[\underset{\sim}{N}(\log \underset{\sim}{p} - \underset{\sim}{\bar{\gamma}}) - \underset{\sim}{y}(\underset{\sim}{\bar{\gamma}})] = \underset{\sim}{0}_{k-m}. \qquad (2.15)$$

Thus, $\underset{\sim}{\bar{\gamma}}$ must satisfy (2.3), and the constraint equations (2.4) and (2.5). This completes the proof.

The system (2.3) - (2.5) contains k + L equations. We can show that L equations of (2.3) are redundant. Without loss of generality, we take the first L rows of $\underset{\sim}{B}_2$ as proportional to the rows of $\underset{\sim}{A}'$, where $\underset{\sim}{A}$ is given by (1.2). Then,

$$\underset{\sim}{A}'[\underset{\sim}{N}(\log \underset{\sim}{p} - \tilde{\underset{\sim}{\gamma}}) - \underset{\sim}{y}(\tilde{\underset{\sim}{\gamma}})]$$

$$= \begin{bmatrix} [\sum_{i \in I_1} n_i(\log p_i - \gamma_i)](1 - \sum_{i \in I_1} \exp \gamma_i) \\ \cdot \\ \cdot \\ \cdot \\ [\sum_{i \in I_L} n_i(\log p_i - \tilde{\gamma}_i)](1 - \sum_{i \in I_L} \exp \tilde{\gamma}_i) \end{bmatrix} = \underset{\sim}{0}_L$$

by (2.5). Computation of $\tilde{\underset{\sim}{\gamma}}$ may be effected through solution of (2.4), (2.5), and the last $(k - m - L)$ equations of (2.3).

2.2 The Approximate Minimum δ Method

Calculation of the minimum δ estimator ecessarily involves solution of equations (2.4), (2.5), and the last $(k - m - L)$ equations of (2.3). As this system is nonlinear, its solution may prove to be a formidable task indeed. In this section, we develop an approximation to the minimum δ estimator, one which is examined as a new estimator. The method is of interest because the approximation, unlike the minimum δ estimator, is relatively easy to compute, and, like the minimum δ estimator, has good asymptotic properties. Furthermore, an approximate minimum δ estimator always exists.

As noted, minimization of δ over $\Gamma(\underset{\sim}{B}_1)$ may prove to be a difficult problem. A somewhat simpler problem is the minimization of δ over $\Gamma_M(\underset{\sim}{B}_1)$, and this is the idea behind the approximate method.

Let $\tilde{\boldsymbol{\gamma}}_a$ denote the approximate minimum d estimator. Roughly, to calculate $\tilde{\boldsymbol{\gamma}}_a$, we first obtain the "projection" of log $\boldsymbol{p}$ in $\Gamma_M(\boldsymbol{B}_1)$ through the minimization of $d(\boldsymbol{\gamma}; \boldsymbol{n})$ over $\Gamma_M(\boldsymbol{B}_1)$. Let $\bar{\boldsymbol{\gamma}}$ denote the result of this minimization. The approximate minimum d estimator $\tilde{\boldsymbol{\gamma}}_a$ is the "projection" of $\bar{\boldsymbol{\gamma}}$ in Γ_S, obtained through minimization of $d^*(\boldsymbol{\gamma}; \bar{\boldsymbol{\gamma}}) = \sum_{i=1}^{k} (\exp \gamma_i - \exp \bar{\gamma}_i)^2/(\exp \bar{\gamma}_i)$ over Γ_S (see Figure A).

The description in the preceding paragraph is oversimplified in that the existence and uniqueness of $\bar{\boldsymbol{\gamma}}$ and $\tilde{\boldsymbol{\gamma}}_a$ are tacitly assumed. In the next section, we show that such points do exist, though they need not be unique.

More formally, let

$$\bar{\Gamma} = \{\bar{\boldsymbol{\gamma}}: \ \bar{\boldsymbol{\gamma}} \text{ minimizes } d \text{ over } \Gamma_M(\boldsymbol{B}_1)\}, \tag{2.16}$$

and

$$\boldsymbol{c}(\boldsymbol{\gamma}) = \boldsymbol{A}' \exp \boldsymbol{\gamma}, \tag{2.17}$$

where $\boldsymbol{A}$ is given by (1.2). An approximate minimum d estimator of $\boldsymbol{\gamma}$ is given by

$$\tilde{\boldsymbol{\gamma}}_a = \bar{\boldsymbol{\gamma}} - \boldsymbol{A}[\log \boldsymbol{c}(\bar{\boldsymbol{\gamma}})], \tag{2.18}$$

where $[\log \underset{\sim}{c}(\bar{\underset{\sim}{\gamma}})]' = [\log c_1(\bar{\underset{\sim}{\gamma}}), \ldots, \log c_L(\bar{\underset{\sim}{\gamma}})]$, and $\bar{\underset{\sim}{\gamma}} \in \Gamma$. We denote the set of points given by (2.18) as Γ_a, and let $\bar{\underset{\sim}{\pi}}_a$ and $\bar{\underset{\sim}{\mu}}_a$ denote the corresponding estimators of $\underset{\sim}{\pi}$ and $\underset{\sim}{\mu}$ respectively.

It is clear that "projection" of $\log \underset{\sim}{p}$ into $\Gamma_M(\underset{\sim}{B}_1)$ lead to the set $\bar{\Gamma}$. In the following lemma, we prove that the "projection" of $\bar{\underset{\sim}{\gamma}} \in \bar{\Gamma}$ in Γ_S, obtained through minimization of $d^*(\underset{\sim}{\gamma}; \bar{\underset{\sim}{\gamma}})$ over Γ_S, is the point $\bar{\underset{\sim}{\gamma}}_a$, given by (2.18). We then prove that any element of Γ_a lies in $\Gamma(\underset{\sim}{B}_1)$, as depicted in Figure A.

Lemma 2.1: The minimum of $d^*(\underset{\sim}{\gamma}; \bar{\underset{\sim}{\gamma}}) = \sum_{i=1}^{k} (\exp \gamma_i - \exp \bar{\gamma}_i)^2/\exp \bar{\gamma}_i$ over Γ_S is attained at $\bar{\underset{\sim}{\gamma}}_a$, where $\bar{\underset{\sim}{\gamma}}_a$ is given by (2.18) for any vector $\bar{\underset{\sim}{\gamma}} \in \bar{\Gamma}$.

Proof: Let $\underset{\sim}{\gamma}_0$ be any element of Γ_S. There exists a k-vector $\underset{\sim}{e}$ such that

$$\exp \gamma_{0i} = e_i + \exp \gamma_{ai}.$$

Since $\underset{\sim}{\gamma}_0$ and $\bar{\underset{\sim}{\gamma}}_a$ are elements of Γ_S, it may be shown that

$$\sum_{i \in I_\ell} e_i = 0, \ \ell = 1, \ldots, L. \qquad (2.19)$$

To prove the lemma, we show that $d^*(\underset{\sim}{\gamma}_0;\ \underset{\sim}{\bar{\gamma}}) \geq d^*(\underset{\sim}{\tilde{\gamma}}_a;\ \underset{\sim}{\gamma})$. Note that

$$d^*(\underset{\sim}{\gamma}_0;\ \underset{\sim}{\bar{\gamma}}) = \sum_{i=1}^{k} (e_i + \exp \gamma_{ai} - \exp \bar{\gamma}_i)^2/\exp \bar{\gamma}_i$$

$$= d^*(\underset{\sim}{\tilde{\gamma}}_a;\ \underset{\sim}{\bar{\gamma}}) + \sum_{i=1}^{k} e_i^2/\exp \bar{\gamma}_i$$

$$- 2 \sum_{\ell=1}^{L} \sum_{i \in I_\ell} e_i [1 - \exp \gamma_{ai}/\exp \bar{\gamma}_i]$$

$$= d^*(\underset{\sim}{\tilde{\gamma}}_a;\ \underset{\sim}{\bar{\gamma}}) + \sum_{i=1}^{k} e_i^2/\exp \bar{\gamma}_i$$

$$- 2 \sum_{\ell=1}^{L} \{1 - [c_1(\underset{\sim}{\bar{\gamma}})]^{-1}\} \sum_{i \in I_\ell} e_i$$

by (2.17) and (2.18). Finally, by (2.19), we have

$$d^*(\underset{\sim}{\gamma}_0;\ \underset{\sim}{\bar{\gamma}}) = d^*(\underset{\sim}{\tilde{\gamma}}_a;\ \underset{\sim}{\bar{\gamma}}) + \sum_{i=1}^{k} e_i^2/\exp \bar{\gamma}_i$$

$$\geq d^*(\underset{\sim}{\tilde{\gamma}}_a;\ \underset{\sim}{\bar{\gamma}})$$

and the proof is complete.

<u>Theorem 2.3</u>: Any $\underset{\sim}{\tilde{\gamma}}_a \in \Gamma_a$, lies in $\Gamma(\underset{\sim}{B}_1)$.

<u>Proof</u>: By (2.17) and (2.18) $\underset{\sim}{\tilde{\gamma}}_a$ lies in Γ_S. To show $\underset{\sim}{\tilde{\gamma}}_a$ lies in $\Gamma_M(\underset{\sim}{B}_1)$, we note that

$$\underset{\sim}{B}_1\tilde{\underset{\sim}{\gamma}}_a = \underset{\sim}{B}_1\bar{\underset{\sim}{\gamma}} - \underset{\sim}{B}_1\underset{\sim}{A}[\log \underset{\sim}{c}(\bar{\underset{\sim}{\gamma}})],$$
$$= \underset{\sim}{0}_m,$$

by (1.1) and (2.16).

3. EXISTENCE AND UNIQUENESS OF THE MINIMUM δ AND APPROXIMATE MINIMUM δ ESTIMATORS

3.1 Introduction

When each cell of the contingency table contains at least one observation, likelihood, minimum δ, and approximate minimum δ estimators may be shown to exist and to be unique. Problems may arise, however, when some cells are empty. In this section, we show that an approximate minimum δ estimator always exists, and give a usable, necessary and sufficient condition for the existence of the minimum δ estimator. The minimum δ estimator is shown to be unique whenever it exists. The approximate minimum δ estimator need not be unique; a necessary and sufficient condition for its uniqueness is given.

Throughout this section, we shall assume that $\underset{\sim}{B}_1$ has less than k - L rows, for when $\underset{\sim}{B}_1$ has k - L rows, $\Gamma(\underset{\sim}{B}_1)$ contains only a single point, and likelihood, minimum δ and approximate minimum δ estimators must exist and be unique.

We make use of the following notation and definition in this section. In analogy to $\Gamma(\underset{\sim}{B}_1)$, for an L-vector $\underset{\sim}{\alpha}$, $\alpha_\ell > 0$, $\ell = 1, \ldots, L$, let

$$\Gamma(\underset{\sim}{B}_1, \underset{\sim}{\alpha}) = \{\underset{\sim}{\gamma}: \underset{\sim}{B}_1\underset{\sim}{\gamma} = \underset{\sim}{0}_m, \sum_{i \in I_\ell} \exp \gamma_i = \alpha_\ell, \ell = 1, \ldots, L\}. \tag{3.1}$$

In this notation $\Gamma(\underset{\sim}{B}_1) = \Gamma(\underset{\sim}{B}_1, \underset{\sim}{1}_L)$.

Definition 3.1: A sequence of k-vectors $\{\underset{\sim}{x}_r\}_{r=1}^{\infty}$ is said to possess the star property if

i) $\{x_{rj}\}_{r=1}^{\infty}$ either converges or diverges properly, $j = 1, \ldots, k$, and

ii) there exists j such that $\lim_r \sup x_{rj} = -\infty$.

3.2 Uniqueness of the Minimum d Estimator

Theorem 3.1: If a minimum of d over $\Gamma(\underset{\sim}{B}_1)$ exists, it is unique.

To prove Theorem 3.1, we require a sequence of three lemmas:

Lemma 3.1: let $\alpha_\ell > 0, \ell = 1, \ldots, L, n_i > 0, i = 1, \ldots, k$, and q be a positive integer. The function,

$$\phi_q(\underset{\sim}{\gamma}; \underset{\sim}{n}) = \sum_{i=1}^{k} n_i(\log p_i - \gamma_i)^q, \tag{3.2}$$

attains a minimum over $\Gamma(\underset{\sim}{B}_1, \underset{\sim}{\alpha})$,

Lemma 3.2: Let $0 < \alpha_\ell \leq 1$. For any ℓ, $\ell = 1, \ldots, L$,

$$\phi_1^{(\ell)}(\gamma; n) = \sum_{i \in I_\ell} n_i(\log p_i - \gamma_i) \geq -N_\ell \log \alpha_\ell \tag{3.3}$$

for $\underset{\sim}{\gamma} \in \Gamma_S(\alpha_\ell) = \{\underset{\sim}{\gamma}: \sum_{i \in I_\ell} \exp \gamma_i = \alpha_\ell\}$,

and

Lemma 3.3: (i) Let $\underset{\sim}{\gamma} \in \Gamma_M(\underset{\sim}{B}_1) \cap \Gamma_S(\alpha_\ell)$, with $0 < \alpha_\ell \leq 1$. Then

$$\begin{aligned} d^{(\ell)}(\underset{\sim}{\gamma}; \underset{\sim}{n}) &= \sum_{i \in I_\ell} n_i(\log p_i - \gamma_i)^2 \\ &\geq d^{(\ell)}(\underset{\sim}{\gamma} - \log \alpha_\ell \underset{\sim}{\Lambda}_\ell; \underset{\sim}{n}), \end{aligned}$$

with strict inequality if $\alpha_\ell < 1$.

(ii) If $\underset{\sim}{\gamma} \in \Gamma(\underset{\sim}{B}_1, \underset{\sim}{\alpha})$, $0 < \alpha_\ell \leq 1$, $\ell = 1, \ldots, L$, with $\alpha_\ell < 1$ for some ℓ, then

$$d(\underset{\sim}{\gamma}; \underset{\sim}{n}) > d(\underset{\sim}{\gamma} - \underset{\sim}{\Lambda} \log \underset{\sim}{\alpha}; \underset{\sim}{n}). \tag{3.4}$$

The proofs of each lemma, and then of Theorem 3.1 are straightforward, and so are omitted.

3.3 Existence and Uniqueness of the Approximate Minimum δ Estimator

The main results of this section are that an approximate minimum δ estimator always exists, and that it is unique if and only if $(\underset{\sim}{B}_2 \underset{\sim}{N} \underset{\sim}{B}_2')$ is nonsingular. These results follow from Lemma 3.4, which describes the nature of $\bar{\Gamma}$ in (2.16), and is stated without proof.

Lemma 3.4: The set $\bar{\Gamma}$ is a nonempty, affine set of dimension $k - m - rk\ (\underset{\sim}{B}_2 \underset{\sim}{N} \underset{\sim}{B}_2')$ and is given by

$$\bar{\Gamma} = \{\bar{\underset{\sim}{\gamma}}: \ \bar{\underset{\sim}{\gamma}} = \underset{\sim}{B}_2[(\underset{\sim}{B}_2 \underset{\sim}{N} \underset{\sim}{B}_2')^- \underset{\sim}{B}_2 \underset{\sim}{N} \log \underset{\sim}{p} \qquad (3.4)$$

$$+ \{I - (\underset{\sim}{B}_2 \underset{\sim}{N} \underset{\sim}{B}_2')^- (\underset{\sim}{B}_2 \underset{\sim}{N} \underset{\sim}{B}_2')\ \underset{\sim}{z}\}],\ \underset{\sim}{z} \in E^{k-m}\},$$

where the notation is such that $\underset{\sim}{A}^-$ is a generalized inverse of $\underset{\sim}{A}$. If $(\underset{\sim}{B}_2 \underset{\sim}{N} \underset{\sim}{B}_2')$ is of full rank, then $\bar{\Gamma}$ is a singleton set. Furthermore, if $\bar{\gamma} \varepsilon \bar{\Gamma}$, $c_\ell(\bar{\underset{\sim}{\gamma}}) \geq 1$, where $c_\ell(\bar{\underset{\sim}{\gamma}})$ is the ℓ^{th} element of $\underset{\sim}{c}(\bar{\underset{\sim}{\gamma}})$ given by (2.17).

Theorem 3.2: The set Γ_a is nonempty. Furthermore Γ_a is a singleton set if and only if $rk(\underset{\sim}{B}_2 \underset{\sim}{N} \underset{\sim}{B}_2') = k-m$.

Proof: The theorem follows immediately from Lemma 3.4 when it is noted that each element $\bar{\underset{\sim}{\gamma}}$ of $\bar{\Gamma}$ leads to a distinct $\tilde{\underset{\sim}{\gamma}} \varepsilon \Gamma_a$, through use of (2.18).

3.4 Existence of the Minimum δ Estimator

The existence of the minimum δ estimator depends upon the space $\bar{\Gamma}$, which was defined in Section 2 and whose properties were noted in the previous section. Let

$$\bar{\Gamma}^* = \{\bar{\underset{\sim}{\gamma}}:\ \bar{\underset{\sim}{\gamma}} \varepsilon \bar{\Gamma},\ \sum_{i \epsilon I_\ell} \exp \bar{\gamma}_i = c_\ell,\ \ell = 1, \ldots, L\}, \qquad (3.5)$$

where

$$c_\ell = \inf_{\bar{\underset{\sim}{\gamma}} \epsilon \bar{\Gamma}} \{c_\ell(\bar{\underset{\sim}{\gamma}})\} = \inf_{\bar{\underset{\sim}{\gamma}} \epsilon \bar{\Gamma}} \{ \sum_{i \epsilon I_\ell} \exp \bar{\gamma}_i\},\ \ell = 1, \ldots, L, \qquad (3.6)$$

In this subsection we show that $\bar{\Gamma}^*$ is either an empty or singleton set, and that a minimum δ estimator exists if and only if $\bar{\Gamma}^*$ is a singleton set.

<u>Lemma 3.5</u>: If $\bar{\Gamma}^*$ is not empty, then it is a single point.

<u>Proof</u>: The proof is by contradiction. Suppose that $\bar{\underset{\sim}{\gamma}}_1$ and $\bar{\underset{\sim}{\gamma}}_2$ are both contained in $\bar{\Gamma}^*$ and that $\bar{\underset{\sim}{\gamma}}_1 \neq \bar{\underset{\sim}{\gamma}}_2$. Since $\bar{\Gamma}$ is an affine set (Lemma 3.4), $\bar{\underset{\sim}{\gamma}}_3 = \frac{1}{2}(\bar{\underset{\sim}{\gamma}}_1 + \bar{\underset{\sim}{\gamma}}_2) \in \bar{\Gamma}$. Furthermore, since the exponential is a convex function, there exists ℓ for which

$$c_\ell = \frac{1}{2} \sum_{j=1}^{2} \sum_{i \in I_\ell} \exp \bar{\gamma}_{ji} > \sum_{i \in I_\ell} \exp \bar{\gamma}_{3i}. \qquad (3.7)$$

But (3.7) contradicts the fact that

$$c_\ell = \inf_{\bar{\gamma} \in \bar{\Gamma}} \sum_{i \in I_\ell} \exp \bar{\gamma}_i = \sum_{i \in I_\ell} \exp \bar{\gamma}_{1i} = \sum_{i \in I_\ell} \exp \bar{\gamma}_{2i},$$

and the lemma is proved.

Let

$$\Gamma_C(\underset{\sim}{B}_1, \underset{\sim}{\alpha}) = \{\underset{\sim}{\gamma}: \ \underset{\sim}{B}_1 \underset{\sim}{\gamma} = \underset{\sim}{0}_m, \ \sum_{i \in I_\ell} \exp \gamma_i \leq \alpha_\ell, \ \ell = 1, \ldots, L\}. \tag{3.8}$$

The following is a technical result needed to prove the main result of this article.

Lemma 3.6: Suppose that δ has a unique minimum over $\Gamma_C(\underset{\sim}{B}_1, \underset{\sim}{\alpha})$. Then, if $\{\underset{\sim}{\gamma}_r^*\}_{r=1}^{\infty}$ is any sequence of points in $\Gamma_C(\underset{\sim}{B}_1, \underset{\sim}{\alpha})$ with the star-property,

$$\lim_r \inf \delta(\underset{\sim}{\gamma}_r^*; \underset{\sim}{n}) = \infty .$$

We now show that the nonemptiness of $\bar{\Gamma}^*$ is necessary and sufficient for the existence of a minimum δ estimator.

Theorem 3.4: A minimum δ estimator exists if and only if $\bar{\Gamma}^*$ is nonempty.

Proof: Suppose that $\bar{\Gamma}^*$ contains the point $\bar{\underset{\sim}{\gamma}}^*$. Then $\bar{\underset{\sim}{\gamma}}^*$ minimizes δ over $\Gamma_C(\underset{\sim}{B}_1, \underset{\sim}{c})$ and is unique by Lemma

3.5. If $\{\underset{\sim}{\gamma}_r\}_{r=1}^{\infty}$ is a sequence of points in $\Gamma(\underset{\sim}{B}_1)$ with the star property, then, by Lemma 3.6,

$$\lim_r \inf (\underset{\sim}{\gamma}_r ; \underset{\sim}{n}) = \infty \quad . \tag{3.9}$$

Note that $\underset{\sim}{a} = -\log(\underset{\sim}{A}\,\underset{\sim}{A}' \,\underset{\sim}{1}_k) \in \Gamma(\underset{\sim}{B}_1)$, and that $\delta(\underset{\sim}{a}\ ;\ \underset{\sim}{n})$ is finite. This, the fact that δ is continuous, and (3.9) imply that a minimum of δ over $\Gamma(\underset{\sim}{B}_1)$ exists. This proves sufficiency.

Now suppose that $\bar{\Gamma}^*$ is empty, and let $\{\underset{\sim}{\gamma}_r\}_{r=1}^{\infty}$ be a sequence of points in $\bar{\Gamma}$ such that

$$\lim_r \sum_{i \in I_\ell} \exp \gamma_{ri} = c_\ell, \ \ell = 1, \ldots, L. \tag{3.10}$$

From the continuity of the exponential function and the definition of c_ℓ in (3.6), such a sequence must exist and, since $\bar{\Gamma}^*$ is empty, must possess the star property.

Let

$$\underset{\sim}{\gamma}_r^* = \underset{\sim}{\gamma}_r - \underset{\sim}{A} \log \underset{\sim}{c}(\underset{\sim}{\gamma}_r), \ r = 1, 2, \ldots, \tag{3.11}$$

and note that $\underset{\sim}{\gamma}_r^* \in \Gamma(\underset{\sim}{B}_1)$, $r = 1, 2, \ldots$. To prove necessity, we need to show that

$$\liminf_r d(\underset{\sim}{\gamma}_r^*;\ \underset{\sim}{n}) < \infty . \qquad (3.12)$$

To prove (3.12), we note that

$$d(\underset{\sim}{\gamma}_r^*;\ \underset{\sim}{n}) = d(\underset{\sim}{\gamma}_r;\ \underset{\sim}{n}) + \sum_{\ell=1}^{L} (\log c_{r\ell})^2 \sum_{i \in I_\ell} n_i + \sum_{\ell=1}^{L} 2 \log c_{r\ell} \sum_{i=I_\ell} n_i(\log p_i - \gamma_i), \qquad (3.13)$$

where $c_{r\ell} = \sum_{i \in I_\ell} \exp \gamma_{ri}$, $\ell = 1, \ldots, L$.

Since

$$\begin{aligned}
\sum_{i \in I_\ell} n_i(\log p_i - \gamma_i) &= \sum_{\{i \in I_\ell : |\log p_i - \gamma_i| \geq 1\}} n_i(\log p_i - \gamma_i) + \sum_{\{i \in I_\ell : |\log p_i - \gamma_i| < 1\}} n_i(\log p_i - \gamma_i) \\
&\leq \sum_{\{i \in I_\ell : |\log p_i - \gamma_i| \geq 1\}} n_i(\log p_i - \gamma_i)^2 + N_\ell \\
&\leq \sum_{i \in I_\ell} n_i(\log p_i - \gamma_i)^2 + N_\ell, \quad \ell = 1, \ldots, L,
\end{aligned}$$

(3.13) implies that

$$\delta(\underset{\sim}{\gamma}_r^*;\ \underset{\sim}{n}) \le \delta(\underset{\sim}{\gamma}_r;\ \underset{\sim}{n}) + \sum_{\ell=1}^{L} N_\ell (\log c_{r\ell})^2$$

$$+ 2 \log C_r [\delta(\underset{\sim}{\gamma}_r;\ \underset{\sim}{n}) + N], \qquad (3.14)$$

where

$$C_r = \max_{\ell=1,\ldots,L} c_{r\ell}. \qquad (3.15)$$

Note that $\delta(\underset{\sim}{\gamma}_r;\ \underset{\sim}{n})$, N, and N_ℓ, $\ell = 1, \ldots, L$ do not depend on r, and that $c_{r\ell}$ and C_r tend to finite limits by (3.10). Thus

$$\lim_r \inf \delta(\underset{\sim}{\gamma}_r^*;\ \underset{\sim}{n}) \le \lim_r \inf\{\delta(\underset{\sim}{\gamma}_r;\ \underset{\sim}{n})$$

$$+ \sum_{\ell=1}^{L} N_\ell (\log c_{r\ell})^2 + 2 \log C_r [\delta(\underset{\sim}{\gamma}_r;\ \underset{\sim}{n}) + N]\}$$

$$< \infty \quad ,$$

which establishes (3.12).

We have exhibited a sequence of vectors $\{\underset{\sim}{\gamma}_r^*\}_{r=1}^{\infty}$ in $\Gamma(\underset{\sim}{B}_1)$, and hence in $\Gamma_C(\underset{\sim}{B}_1, \underset{\sim}{1}_L)$, which violates the conclusion of Lemma 3.6, and hence implies that no unique minimum of δ over $\Gamma_C(\underset{\sim}{B}_1, \underset{\sim}{1}_L)$ exists. Thus, either no minimum exists or the minimum is not unique.

Suppose two distinct vectors $\underset{\sim}{\gamma}_1$ and $\underset{\sim}{\gamma}_2$ which minimize δ over $\Gamma_C(\underset{\sim}{B}_1, \underset{\sim}{1}_L)$ exist. By part (ii) of Lemma

3.3, $\underset{\sim}{\gamma}_1$ and $\underset{\sim}{\gamma}_2$ must lie in $\Gamma(\underset{\sim}{B}_1)$. But this contradicts Theorem 3.1, which states that, if a minimum of d over $\Gamma(\underset{\sim}{B}_1)$ exists it is unique. Thus no minimum exists. This proves necessity, and the proof of the theorem is complete.

The following corollary is useful and follows from Theorems 3.2 and 3.4.

Corollary 3.4: If $(\underset{\sim}{B}_2 \, \underset{\sim}{N} \, \underset{\sim}{B}_2')$ is of rank k-m, then a minimum d estimator exists.

4. ASYMPTOTIC PROPERTIES

As the sample sizes become large, since $\pi_i > 0$, the probability that each cell in the contingency table contains at least one observation tends to one. Therefore, the probability that the defined estimators exist and are unique tends to one by Corollary 3.4 and Theorem 3.2. Thus we assume that the various estimators exist and are unique and consider their asymptotic properties. Further, we assume that there exists an L-vector $\underset{\sim}{e}$, $e_\ell > 0$, $\ell = 1, \ldots, L$, such that

$$\lim_{N, N_\ell \to \infty} N_\ell / N = e_\ell ,$$

where N_ℓ denotes the sample size for population ℓ, $\ell = 1, \ldots, L$, and $N = \sum_{\ell=1}^{L} N_\ell$ denotes total sample size. Two main results are given in this section.

The first result is Theorem 4.1. Two sequences of random k-vectors, $\{\underset{\sim}{Y}_N\}_{N=1}^{\infty}$ and $\{\underset{\sim}{Y}_N^*\}_{N=1}^{\infty}$, are said to be a_N-limit equivalent in probability (a_N - l.e.p.) if

$$a_N(\underset{\sim}{Y}_N - \underset{\sim}{Y}_N^*) \xrightarrow{p} \underset{\sim}{0}_k. \tag{4.1}$$

Theorem 4.1: Any pair of the maximum likelihood estimator $\hat{\underset{\sim}{\gamma}}$, the minimum δ estimator $\tilde{\underset{\sim}{\gamma}}$, the approximation $\tilde{\underset{\sim}{\gamma}}_a$, and $\bar{\underset{\sim}{\gamma}}$ are $\sqrt{N}$ - l.e.p. .

Proof: We prove that $\hat{\underset{\sim}{\gamma}}$ and $\bar{\underset{\sim}{\gamma}}$ are $\sqrt{N}$ - l.e.p. to give an indication of the method of proof of the various pairwise results.

Likelihood estimation is reviewed in Section 1. As a slight generalization of (1.11), we have

$$\underset{\sim}{B}_1 \hat{\underset{\sim}{\gamma}} = \underset{\sim}{0}_m, \tag{4.2}$$

and

$$\underset{\sim}{B}_2 \underset{\sim}{E} (\underset{\sim}{p} - \exp \hat{\underset{\sim}{\gamma}}) = \underset{\sim}{0}_{k-m} \tag{4.3}$$

where $\underset{\sim}{E}$ denotes the diagonal matrix whose i^{th} element equals e_ℓ, $i \in I_\ell$, $\ell=1, \ldots, L$. From the definition of $\bar{\underset{\sim}{\gamma}}$ in Section

2, the minimization procedure yields

$$\underset{\sim}{B}_1 \bar{\underset{\sim}{\gamma}} = \underset{\sim}{0}_m \tag{4.4}$$

and

$$\underset{\sim}{B}_2 \underset{\sim}{N}(\log \underset{\sim}{p} - \bar{\underset{\sim}{\gamma}}) = \underset{\sim}{0}_{k-m}. \tag{4.5}$$

Replacement, in (4.3), of $\exp \hat{\gamma}_i$, $i = 1, \ldots, k$, by its expansion about $\log p_i$ yields

$$\underset{\sim}{B}_2 \underset{\sim}{E} [p_i - p_i - p_i(\hat{\gamma}_i - \log p_i) - \tfrac{1}{2}(\exp \psi_i)(\hat{\gamma}_i - \log p_i)^2]$$

$$= -N^{-1} \{\underset{\sim}{B}_2 \underset{\sim}{N}(\hat{\underset{\sim}{\gamma}} - \log \underset{\sim}{p}) + \tfrac{1}{2}\underset{\sim}{N} \underset{\sim}{B}_2 \underset{\sim}{E} \underset{\sim}{\Psi} \begin{bmatrix} (\hat{\gamma}_1 - \log p_1)^2 \\ \cdots \\ (\hat{\gamma}_k - \log p_k)^2 \end{bmatrix}\}$$

$$= \underset{\sim}{0}_{k-m}, \tag{4.6}$$

where $\underset{\sim}{\Psi}$ is the diagonal matrix with elements $\exp \psi_1, \ldots, \exp \underset{\sim}{\psi}_k$, and $\underset{\sim}{\psi}' = (\psi_1, \ldots, \psi_k)$ is such that ψ_i lies between $\log p_i$ and $\hat{\gamma}_i$, $i = 1, \ldots, k$.

Subtraction of N times the expression in (4.6) from (4.5), and subtraction of (4.4) from (4.2) yield

$$\underset{\sim}{B}_1(\hat{\underset{\sim}{\gamma}} - \bar{\underset{\sim}{\gamma}}) = \underset{\sim}{0}_m \qquad (4.7)$$

and

$$\{\underset{\sim}{B}_2\ \underset{\sim}{N}(\hat{\underset{\sim}{\gamma}} - \bar{\underset{\sim}{\gamma}}) + \tfrac{1}{2}\underset{\sim}{N}\ \underset{\sim}{B}_2\ \underset{\sim}{E}_N\ \underset{\sim}{\Psi} \begin{vmatrix} (\hat{\gamma}_1 - \log p_1)^2 \\ \cdot \\ \cdot \\ \cdot \\ \hat{\gamma}_k - \log p_k)^2 \end{vmatrix} \} = 0_{k-m}, \qquad (4.8)$$

Slutsky's Theorem (Serfling, 1980, p. 19) implies that

$$\{N^{-1/2}\ \underset{\sim}{B}_2\ \underset{\sim}{N}(\hat{\underset{\sim}{\gamma}} - \bar{\underset{\sim}{\gamma}}) - \sqrt{N}\ \underset{\sim}{B}_2\ \underset{\sim}{\Pi}\ \underset{\sim}{E}(\hat{\underset{\sim}{\gamma}} - \bar{\underset{\sim}{\gamma}})\} \xrightarrow{p} \underset{\sim}{0}_k, \qquad (4.9)$$

where $\underset{\sim}{\Pi}$ denotes the diagonal matrix with elements $\pi_1, \ldots, \pi_k$, since $n_i = p_i N_\ell$, $i \in I_\ell$, and $p_i \xrightarrow{p} \pi_i$.

To complete proof of this part of the theorem, it remains to show that

$$\sqrt{N}\,\underset{\sim}{B}_2\,\underset{\sim}{E}\,\underset{\sim}{\psi}\begin{bmatrix}(\hat{\gamma}_1 - \log p_1)^2\\ \cdot\\ \cdot\\ \cdot\\ (\hat{\gamma}_k - \log p_k)^2\end{bmatrix} \xrightarrow{p} \underset{\sim}{0}_{k-m}. \qquad (4.10)$$

Since $\underset{\sim}{\hat{\gamma}}$ and $\log \underset{\sim}{p}$ are consistent estimators of $\underset{\sim}{\gamma}$, $\underset{\sim}{\psi} \xrightarrow{p} \underset{\sim}{\gamma}$, and, since $\sqrt{N}(\hat{\gamma}_i - \gamma_i)$ and $\sqrt{N}(\log p_i - \gamma_i)$ have limiting distributions, Slutsky's Theorem determines that

$$\sqrt{N}\begin{vmatrix}(\hat{\gamma}_1 - \log p_1)^2\\ \cdot\\ \cdot\\ \cdot\\ (\hat{\gamma}_k - \log p_k)^2\end{vmatrix} \xrightarrow{p} \underset{\sim}{0}_k . \qquad (4.11)$$

Finally, (4.7)-(4.10) imply that

$$\sqrt{N}\begin{bmatrix}\underset{\sim}{B}_2\,\underset{\sim}{\Pi}\,\underset{\sim}{E}\\ \\ \underset{\sim}{E}_1\end{bmatrix}(\underset{\sim}{\hat{\gamma}} - \underset{\sim}{\bar{\gamma}}) \xrightarrow{p} \underset{\sim}{0}_k, \qquad (4.12)$$

and, since $\begin{bmatrix}\underset{\sim}{B}_2\,\underset{\sim}{\Pi}\,\underset{\sim}{E}\\ \\ \underset{\sim}{B}_1\end{bmatrix}$ is of full rank,

$$N(\hat{\underset{\sim}{\gamma}} - \bar{\underset{\sim}{\gamma}}) \xrightarrow{p} \underset{\sim}{0}_k,$$

and the proof is complete.

The second theorem of this section follows.

Theorem 4.2:

$$d(\underset{\sim}{\gamma}^*;\ \underset{\sim}{n}) - X^2(\underset{\sim}{\gamma}^*;\ \underset{\sim}{n}) \xrightarrow{P} 0, \qquad (4.13)$$

where the function

$$X^2(\underset{\sim}{\gamma};\ \underset{\sim}{n}) = \sum_{\ell=1}^{L} \sum_{i \in I_\ell} N_\ell\ (p_i - \exp \gamma_i)^2/\exp \gamma_i$$

is the well-known Pearson function, and $\underset{\sim}{\gamma}^*$ denotes any of the likelihood, minimum d, or approximate minimum d estimators.

This result is proven by expansion of $\exp \gamma_i^*$ in the numerator of $X^2(\underset{\sim}{\gamma}^*;\ \underset{\sim}{n})$ about $\log p_i$.

Theorem 4.2 allows us to test the null hypothesis,

$$H_o: \begin{bmatrix} \underset{\sim}{B}_1 \\ \underset{\sim}{B}_1^* \end{bmatrix} \underset{\sim}{\gamma} = \underset{\sim}{0}_{m+m^*},$$

where $\underset{\sim}{B}_1^*$ is an $m^* \times k$ matrix with orthonormal rows such that $\underset{\sim}{B}_1^* \underset{\sim}{A} = \underset{\sim}{0}_{m^* \times L}$, against the alternative hypothesis,

$$H_a: \underset{\sim}{B}_1 \underset{\sim}{\gamma} = \underset{\sim}{0}_m.$$

Define $T = \delta(\underset{\sim}{\gamma}^*_{Ho}; \underset{\sim}{n}) - \delta(\underset{\sim}{\gamma}^*_{Ha}; \underset{\sim}{n})$,

where $\underset{\sim}{\gamma}^*_{H0}$ and $\underset{\sim}{\gamma}^*_{Ha}$ are estimates of $\underset{\sim}{\gamma}$ under H_0 and H_a respectively. The limiting distribution of test statistics based on X^2, and, in view of Theorem 4.2, of T are given by Mitra (1958) and Diamond (1963) for various m^* and "local alternatives". In particular, under H_0, T has a limiting chi-squared distribution with m^* degrees of freedom.

5. CONCLUDING REMARKS

The emphasis in this paper has been on the basic properties of the minimum δ and approximate minimum δ estimators. Their existence properties have been considered and they have been shown to have large sample properties equivalent to those of likelihood estimators. The minimum δ and approximate minimum δ methods lead to asymptotic chi-squared tests analogous to those of the Pearson goodness of fit statistic.

A second manuscript will deal with the application of log-linear models and the new estimators to the classification problem of Martin and Bradley (1972). The existence results and test procedures will be illustrated there, and extensions to the problem considered by Martin and Bradley (1972) given.

ACKNOWLEDGEMENTS

This article is based on part of a doctoral dissertation, Redman (1980), prepared under the supervision of Ralph A. Bradley. His assistance is greatly appreciated and gratefully acknowledged. He was the Principal Investigator of the referenced Office of Naval Research Contracts.

We acknowledge also the assistance of Frederick W. Leysieffer who directed the referenced training grant supported by the National Institution of Environmental Health Sciences.

REFERENCES

Andersen, A. H. (1974) "Multidimensional Contingency Tables", Scandinavian Journal of Statistics, 1, 115-127.

Apostol, T. M. (1957), Mathematical Analysis, Reading, Mass.: Addison-Wesley.

Bartlett, M. S. (1935), "Contingency Table Interactions", Journal of the Royal Statistical Society, Ser. B, 2, 248-252.

Birch, M. W. (1963), "Maximum Likelihood in Three-Way Contingency Tables," Journal of the Royal Statistical Society, Ser. B, 25, 220-232.

Bishop, Y. M., Feinberg, S. E., and Holland, P. W. (1975), Discrete Multivariate Analysis: Theory and Practice, Cambridge: The MIT Press.

Diamond, E. L. (1963), "The Limiting Power of Categorical Data Chi-Square Tests Analogous to Normal Analysis of Variance," The Annals of Mathematical Statistics 34, 1432-1441.

Ferguson, T. (1958), "A Method of Generating Best Asymptotically Normal Estimates with Application to Estimation of Bacteria Densities," The Annals of Mathematical Statistics, 29, 1046-1062.

Grizzle, J. E., Starmer, C. F., and Koch, G. G., (1969), "Analysis of Categorical Data by Linear Models," Biometrics, 25, 489-501.

Haberman, S. J. (1973), "Log-linear Models for Frequency Data: Sufficient Statistics and Likelihood Equations," The Annals of Statistics, 1, 617-632.

Haberman, S. J. (1974), The Analysis of Frequency Data, Chicago: The University of Chicago Press.

Koch, G. G., Landis, J. R., Freeman, J. L., Freeman, D. H. Jr., and Lehnen, R. G. (1977), "A General Methodology for the Analysis of Experiments with Repeated Measurement of Categorical Data," Biometrics, 33, 133-158.

Martin, D. C., and Bradley, R. A. (1972), "Probability Models, Estimation, and Classification for Multivariate Dichotomous Populations," Biometrics, 28, 203-222.

Mitra, S. K. (1958), "On the Limiting Power Function of the Frequency Chi-Square Test," The Annals of Mathematical Statistics, 29, 1221-1230.

Neyman, J. (1949), "Contribution to the Theory of the X^2 test," Proc. Berkeley Symposium on Math. Stat. and Prob., Berkeley: University of California Press.

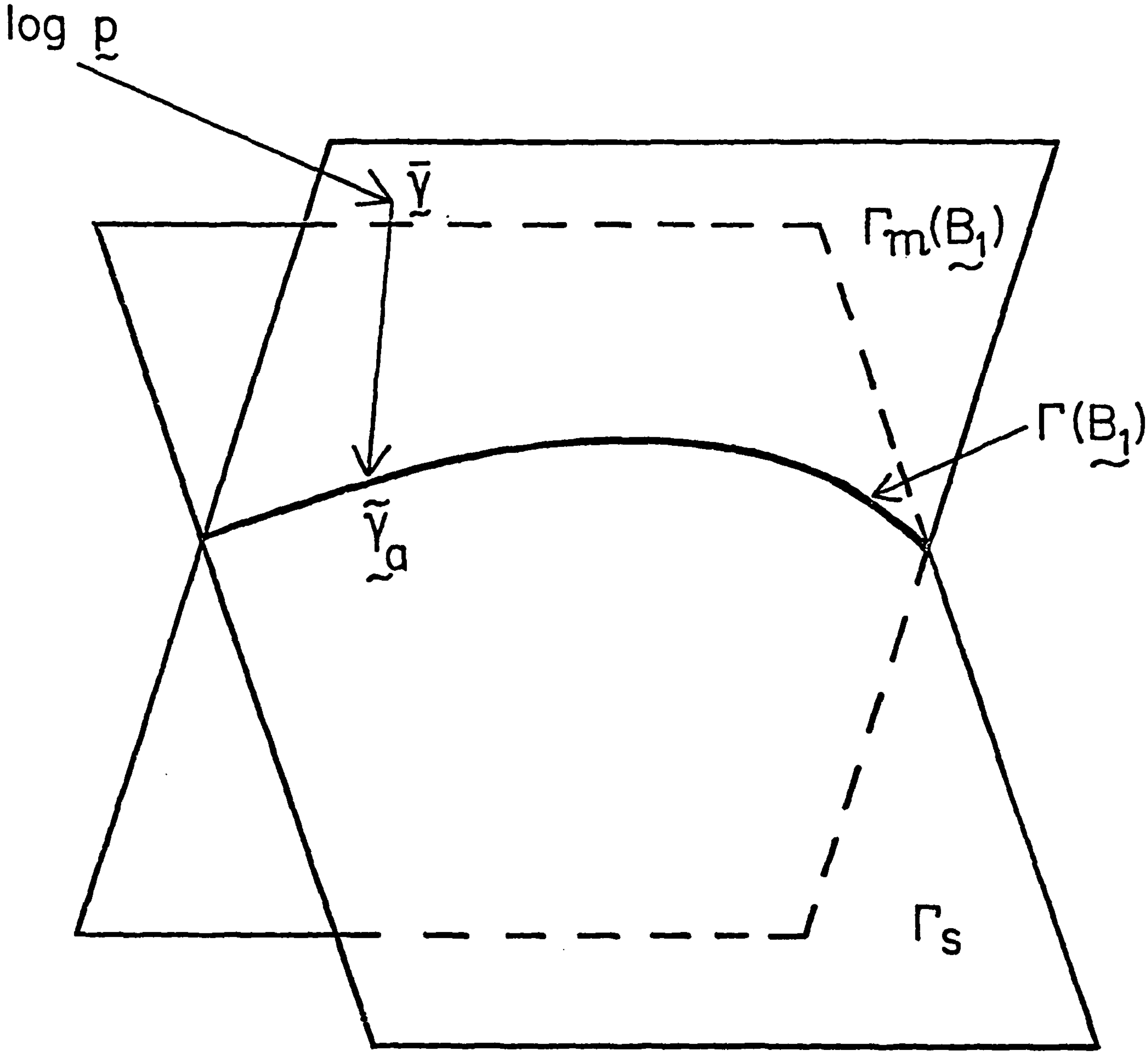
Redman, T. C. (1980), Minimum δ Estimation for Log-Linear Models, unpublished Ph.D. dissertation, The Florida State University, Tallahassee, Florida.

Roy, S. N. and Kastenbaum, M. A. (1956), "On the Hypothesis of No Interaction in a Multi-Way Contingency Table," The Annals of Mathematical Statistics, 27, 749-757.

Serfling, R. J. (1980), Approximation Theorems of Mathematical Statistics, New York: John Wiley & Sons, Inc.

Taylor, W. F. (1953), "Distance Functions and Regular Best Asymptotically Normal Estimates," The Annals of Mathematical Statistics, 24, 85-92.

Williamson, R. E., Crowell, R. H., and Trotter, H. F. (1972), Calculus of Vector Functions, London: Prentice-Hall International.



A. Geometrical interpretation of the approximate minimum δ estimator. The "projection" of log p into $\Gamma_M(B_1)$ is $\bar{\gamma}$. The "projection" of $\bar{\gamma}$ into Γ_S is $\tilde{\gamma}_a$.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1. REPORT NUMBER FSU Report No. M562 ONR Report No. 151	2. GOVT. ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and subtitle) New Estimation Methods for Log-Linear Models	5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER FSU Report M562
7. AUTHOR(s) Thomas C. Redman	8. CONTRACT OR GRANT NUMBER(s) ONR N00014-76-C-0394 N00014-80-C-0093
9. PERFORMING ORGANIZATION NAME AND ADDRESS The Florida State University Department of Statistics Tallahassee, Florida 32306	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Statistics and Probability Program Office of Naval Research Arlington, Virginia 22217	12. REPORT DATE August, 1981 13. NUMBER OF PAGES 36
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in 20, if different from report).

18. SUPPLEMENTARY NOTES

19. KEY WORDS

Minimum δ estimation, approximate minimum δ estimation, existence, limiting distribution.

20. ABSTRACT

Two new methods for estimation of parameters in log-linear models are proposed and their properties considered in this article. In particular, conditions for the existence of the new estimators, more easily checked than those available for likelihood estimators, are derived, and the new estimators are shown to posses appropriate asymptotic properties.

END

DATE
FILMED
11-81

DTIC